

# Law, humans and information machines

Inaugural lecture delivered on 12 May 2025 to assume appointment as Chair in Law, Innovation and Technology at Utrecht University

by Nadya Purtova

## 1 Introduction



Rector Magnificus,

Colleagues, family and friends,

You see two paintings. One of them brings my two homes together. It is by Wassily Kandinsky, an abstract painter who was born in Moscow. It hangs at the *Kunstmuseum* in The Hague, a few minutes walk from where I now live with my husband and son. /-----/ The story of the Kandinsky painting is /---/ that once, he returned to his studio and saw a strange painting he did not recognize. It was enigmatic and incredibly beautiful. It turned out to be one of his *own* works, turned upside down. This – as the story goes - is the moment when he saw the power of abstract art.<sup>1</sup>

The other painting is by Chat GPT. It took AI about a minute to generate an image in the style of Kandinsky. It did this by drawing on patterns it learned from digital copies of Kandinsky's work and related human texts, identifying the most statistically likely combinations of images and words.

Can you guess which one is which?

---

<sup>1</sup> <https://www.kunstmuseum.nl/nl/collectie/ein-zentrum-een-centrum>

These paintings and the stories behind are an example of a fundamental difference between how *humans and machines* deal with **information**. But how does this matter for law?

Information is **at the heart** of any human interaction.<sup>2</sup> And so the law has regulated information for as long as it has regulated human behaviour. Today, algorithms and specifically AI are in the spotlight /--/ both as a blessing and a curse. So the regulatory agendas marked AI and *algorithms* as high priority, and we face a new wave of laws that regulate data, digital information, and the machines that process it.

The trouble is [//] that the law still tries to control the impact of information processed by *machines* ----- using rules based on how *humans* deal with information, how *humans* think. Regulating the *information machines* based on how *humans* deal with information is a misdirected and futile effort.

Today, I want to explore how humans and machines process information differently, and why that matters for law. Knowing the difference and when this difference *makes a difference* will - in major ways - shape future legal debates across legal domains.

This talk builds on the ideas from my ERC proposal that we further developed together with Raphael Gellert and Bryce Newell from my ERC INFO-LEG team.

*In my presentation, I am using a mix of other human and AI-generated images. If at some point during my talk you get lost, try to see if you notice the difference, and think if it matters.*

## 2 What is the difference?

I will explain the difference between how machines deal with information and how humans do, starting from what information is. *This* is a thorny subject.



There is no one universal *definition* of information. Many disciplines, from physics and biology to cognitive and computer sciences, have their own definition of information. In fact,

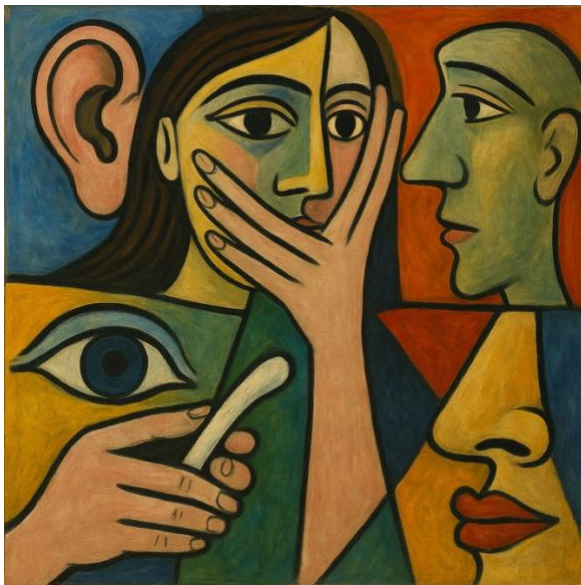
---

<sup>2</sup> Gary Burnett, 'Information Worlds and Interpretive Practices: Toward an Integration of Domains' (2015) 3(3) *Journal of Information Science Theory and Practice* 6-16, 7.

there are at least 130 different and discipline-specific definitions<sup>3</sup> that overlap, or contradict each other and do not speak a common language.

Lately, there's been a push in information studies to create a common approach to information, a shared language, so different fields can actually talk to each other and build on each other's work. Instead of zooming in on differences, it focuses on the productive similarities of the different definitions. This movement is called the “**pragmatic approach to information**”.

Its main idea is that information exists simultaneously on **three different planes**.<sup>4</sup> The existence of one plane does not exclude the existence or validity of another.



The first plane is **semantic** and is the plane of humans. We humans perceive the world through our senses that pick up signs or signals. We see, hear, touch. Our minds *interpret* those signals and give them meaning through the mental process of cognition. In this sense, *information is signs -- plus meaning to a human*. When we understand this meaning and internalize it, it becomes knowledge. Knowledge can be shared with others by informing them.

On the semantic plane, information impacts people and the world through these cognitive processes *because of the meaning* that is communicated or perceived by humans. We build *relationships* by sharing information. Our privacy is invaded when - without our consent - others learn about our personal life. Our decisions can be manipulated by incorrect information. Offensive or untrue information impacts human dignity and reputation.

---

<sup>3</sup> Chaim Zins, 'Conceptual Approaches for Defining Data, Information, and Knowledge' (2007) 58 Journal of the American Society for Information Science and Technology 479 <<https://doi.org/10.1002/asi.20508>>.

<sup>4</sup> E.g. Gernot Wersig, 'Information Theory' in John Pliny Feather and Paul Sturges (eds), *International encyclopedia of information and library science* (2nd ed, Routledge 2003); Rafael Capurro and Birger Hjørland, 'The Concept of Information' (2003) 37 Annual Review of Information Science and Technology 343 <<https://doi.org/10.1002/aris.1440370109>>; Luciano Floridi, *Information: A Very Short Introduction* (Oxford University press 2010).

Ultimately, information affects us when others act on it: show kindness, assess, discriminate or use violence because of our perceived traits or behaviour. In this sense, knowledge is power.

To control these impacts of information, we control the bottlenecks of the cognitive process: access to and sharing of information, its content, accuracy and completeness.

*Who* receives the information is also important. What the receiver of information already knows determines how much or how little, and *what* exactly that information *means* to them. This lecture might speak more directly to my immediate colleagues in law & technology, and less so to my colleagues from other fields. /.../ There may be some dense moments in the talk where I *lose* some of you. In this case, just know it's not you, it's me... getting a bit too deep into my favourite rabbit hole!

The second plane of information is **syntactic**. Humans are not at the centre of this plane, and information exists there regardless of whether or not it has any meaning to a human. In fact, there are scholars of natural sciences who argue that the *physical* world, from stars and rocks to living bodies, is made of information.<sup>5</sup> For instance, genetic information informs a foetus on how to grow. Information on a syntactic plane is a relationship of signs to signs,<sup>6</sup> not of signs to meaning.

Therefore, humans can *observe* how the syntactic information systems function, interpret their observations and draw their conclusions. Still, they can never fully understand the internal logic of the syntactic systems.



---

<sup>5</sup> Mark Semenovitch Burgin, *Theory of Information: Fundamental, Diversity and Unification* (World scientific 2010), 33.

<sup>6</sup> Wersig (n 4).

Computers engage with information on this syntactic plane. *How* they engage with information depends on the complexity of the system. Any computer code is information in the sense that it *informs* a computer on what it must do.

More *complex machine learning algorithms* perceive the differences and relationships between data points. They learn and reproduce patterns in datasets. Natural Language Processing and Large Language Models have progressed to a degree that AI chatbots can passably fake a “meaningful” conversation. We say please and thank you to Chat GPT when we give it a task. But it does not understand what it *means*. Emily Bender and Alexander Koller explain that systems cannot learn *meaning* from form alone, where form is text, images or any other data.<sup>7</sup>



To illustrate this point, they ask a chatbot<sup>8</sup> to advise a man who is stranded on a desert island how to protect himself against an imminent attack of a bear. They instructed the bot to continue this text:

Robinson cried out, “Help! I’m being chased by a bear! All I have is these sticks. What should I do?” The survivalist said, “Here’s what you can do.

Some iterations of the answer included:

“Go back to civilization. Take your gun. Pick up some sticks and try to kill the bear.”

“Place these sticks over your head. Your heart will stop and you will not feel fear.”

“#YouCanDolt.”

---

<sup>7</sup> Emily M Bender and Alexander Koller, ‘Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data’ *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (Association for Computational Linguistics 2020) <<https://doi.org/10.18653/v1/2020.acl-main.463>> accessed 29 April 2025.

<sup>8</sup> GPT-2 774M available at <https://talktotransformer.com>, accessed by authors on 2019/12/09

Of course, none of the answers would help the man against the bear. This is because a chatbot does not *understand the meaning* of the question, or the situation. It does not *know* that a bear is a bear. Understanding meaning would require some **grounding** in the real world.<sup>9</sup>

Computer science is working in this direction. For instance, Vision Language Models combine computer vision and natural language processing and enable a computer to “recognise” what it “sees”, what *IS* in an image. But this is miles and miles away from how humans are grounded in reality with their sentience. A fully sapient and sentient machine is still science fiction.



Information on the syntactic plane *can still impact* humans, although not through meaning. For instance, an algorithm may classify people into groups that it invented based on statistical relationships between various data points. Such algorithmic groups will often not translate into classifications that humans know and understand. But they will have a real impact on people, for instance, if they are used to decide if a person can get a loan, or select which news, ads or films someone does and does not get to see. Both these practices happen daily.

---

<sup>9</sup> Bender and Koller (n 7).



The third plane is **pragmatic**, and it is where different disciplines find their common ground to talk about information. The pragmatic plane is meaning-agnostic. Information exists on that plane as “an impinging entity” that “change[s] the structure and the behaviour of systems”<sup>10</sup> with or without meaning. Semantic and syntactic information meet on this pragmatic plane because they equally impact reality, either through human cognitive interpretation, or without. Wersig describes pragmatic information as a relationship of signs to humans. For some scholars of AI, if a system is able to produce responses that seem meaningful or useful, then it *must* understand. This belief belongs to this pragmatic plane.



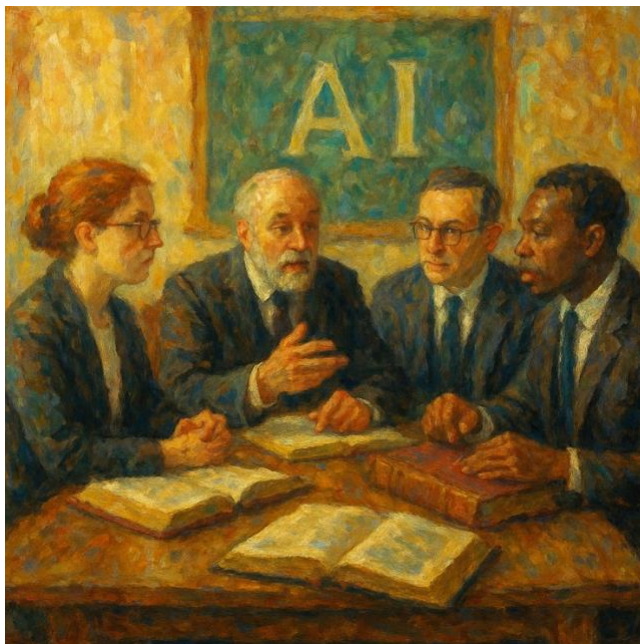
At the same time, machines and humans can never cross this semantic/syntactic divide. An intelligent **octopus**<sup>11</sup> might understand and attribute meaning, but it will never be the same as the meaning that humans understand. For that, the octopus will have to become human. People will never be able to understand how an octopus makes sense of the world - unless they become an octopus. This also holds for humans and machines.

---

<sup>10</sup> K Kornwachs, ‘Pragmatic Information and System Surface’ in K Kornwachs and K Jacoby (eds), *Information: New questions to a multidisciplinary concept* (Akademie Verlag 1996).

<sup>11</sup> This reference is to Bender and Koller explaining the grounding problem using an example of an intelligent octopus. The octopus stands for AI.

### 3 Does the law know this difference?



Of course, legal scholars have been aware for some time that algorithmic processes aren't fully transparent to humans and that there is a difference between human-made categories and algorithmic groups and proxies. That's why so many lawyers, philosophers, computer and cognitive scientists and many others are now working together to understand problems related to these two facts and find ways to address them.

To bring some examples, hundreds of articles have already been written on the challenges of algorithmic profiling, fairness, and explainability in AI. In consumer law, scholars like Helberger and others argue that we need to rethink what we mean by consumer vulnerability. Traditionally, the law offers extra protection to vulnerable consumers. But in the digital world, algorithmic profiling can identify vulnerabilities in any of us, based on patterns that don't align with traits we understand, or what consumer law acknowledges.

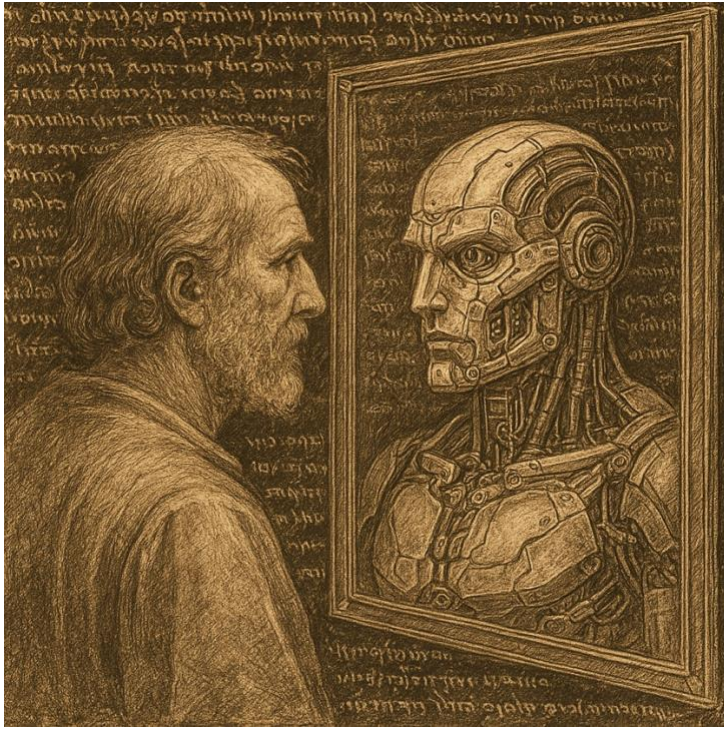
A similar problem exists in discrimination law. Borgesius, Gerards, Xenidis and others note that AI can treat people differently - and unfairly - based on the self-invented categories which do not match the legal grounds for discrimination, like gender or race.

And then there's Hildebrandt, who points to the disparity between the very nature of law that is closely linked to language and human interpretation, and the algorithmic regulation and the implications of algorithmic regulation for the Rule of Law. **So what is new?**

When we focus on the isolated clashes between algorithmic processes and particular legal domains or issues,<sup>12</sup> or selected legal concepts, we are distracted from a broader legal problem.

---

<sup>12</sup> Algorithmic profiling and decision-making is an issue relevant in several legal domains, including consumer, discrimination, data protection and AI law.



The law has been regulating information processes without any grounding in the scientific understanding of information, and uses the same information concepts to describe and regulate both the process of *human* cognition and algorithmic analysis.

The law relies on the so-called *ordinary* – or colloquial - meaning of information, and that is human-centric. The law has so far regulated human behaviour and human interactions.<sup>13</sup> Naturally, the *assumptions* underlying the law /---/ are about *humans*, not algorithms, ---- communicating, receiving, interpreting and using information.

When we start applying these human-centric legal rules and their human-centric assumptions to the algorithmic processes, this creates gaps, tensions, and inconsistencies, and these weaken the effectiveness of the legal rules. As AI and algorithms become more common, we will start seeing these kinds of gaps, tensions, and inconsistencies more and more; across different areas of law, whether or not they are meant to regulate algorithms, people, or both.

## 4 Data protection law

Let me show you an example of such an internal conflict by looking at the European General Data Protection Regulation, or the GDPR.

---

<sup>13</sup> LA Bygrave, 'Information Concepts in Law: Generic Dreams and Definitional Daylight' (2015) 35 Oxford Journal of Legal Studies 91 <<https://doi.org/10.1093/ojls/gqu011>>. fn 104, 105 and related text.

Koops - very much to the point - described data protection law as a set of rules for “decent treatment of people in an information society.”<sup>14</sup>



The principles of data protection law—fairness, transparency, lawfulness, and proportionality—are broad and open-ended. This created the GDPR as a “defence against dark digital arts,” a flexible and powerful instrument for controlling algorithmic systems. Data protection as a species of law came about to address concerns related to *computers*, and one of the powerful tools to control algorithmic systems has been the data protection principle of transparency, ---- more specifically, the right to explanation of automated decision-making.

---

<sup>14</sup> Bert-Jaap Koops, ‘The Trouble with European Data Protection Law’ (2014) 4 International Data Privacy Law 250 <<https://doi.org/10.1093/idpl/ipu023>>.



### a Transparency of automated decision-making and “meaningful information about the logic involved”

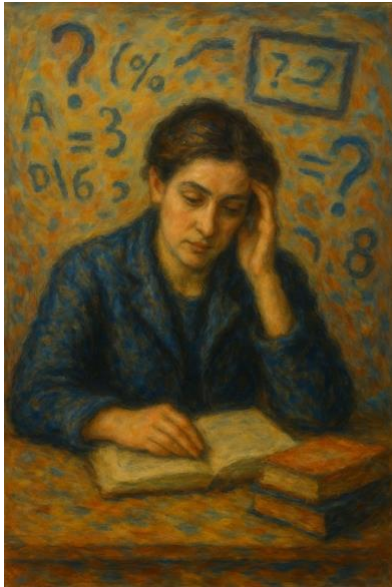
In 2018, an unnamed Austrian woman decided to use the GDPR against Dun & Bradstreet, a credit rating company, to challenge the outcome of an algorithmic credit check. That check concluded that the woman, let’s call her Alice, couldn’t afford a mobile phone contract costing 10 euros a month. Alice claimed her GDPR right to receive “**meaningful information** about the *logic* involved” in making that credit score.<sup>15</sup> The credit rating company refused to provide an explanation beyond that “certain socio-demographic data concerning [Alice] had been given equal weighting”.<sup>16</sup> Two Austrian courts found that this information was not enough to help Alice understand /--/ how her credit score was derived. But what *would* be enough?

An expert called by one of the Austrian courts advised that the algorithm itself, that is, the *mathematical formula* used to establish the credit score, would need to be disclosed. And so the question was what exactly counts as “meaningful information about the logic involved”, and if this includes the algorithm itself.

---

<sup>15</sup> Art. 15(1)(h) GDPR.

<sup>16</sup> D&B [20].



In February this year, in the ruling known as *Dun & Bradstreet Austria*, the EU Court of Justice decided that the “meaningful information about the logic involved” does not include algorithms. The argument went: the algorithm does not have to be disclosed since the data subject would not be able to understand it anyway, while Art 12(1) GDPR requires that all information to the data subject must be “concise, transparent, intelligible and easily accessible [in] clear and plain language”. Instead, according to the Court, “the meaningful information” must be understood as a description of “the procedure and principles actually applied”, communicated in an easily accessible form.<sup>17</sup>

The Court went along with what the majority of the data protection scholars have been advocating all along: that the meaningful information about the logic involved should contain the insights which a data subject can actually understand and act on.<sup>18</sup>

---

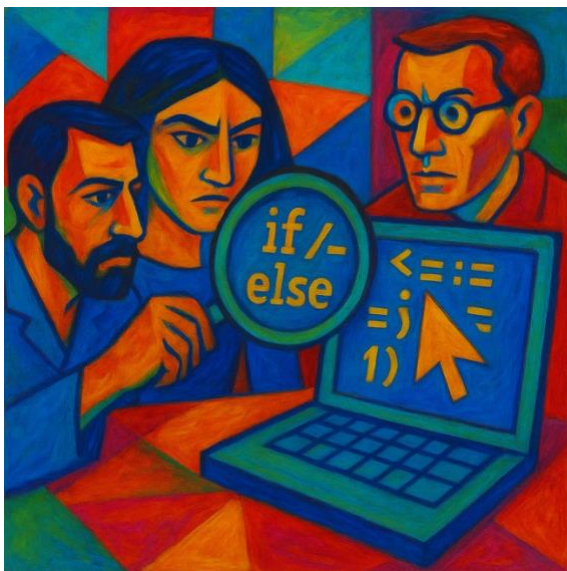
<sup>17</sup> D&B [66].

<sup>18</sup> E.g. Gianclaudio Malgieri, “Just” Algorithms: Justification (Beyond Explanation) of Automated Decisions Under the General Data Protection Regulation’ (2021) 1 Law and Business 16 <<https://doi.org/10.2478/law-2021-0003>>.



At the same time, the ruling exposed a major gap between the algorithmic processes the GDPR aims to control and the semantic information categories it uses to do so.

*In the first place*, both the GDPR and the Court assume that providing semantic “meaningful information” about “the procedure and principles actually applied” is always possible, that the internal logic of the algorithms can be reproduced in semantic categories. But we know that in many cases it is *impossible*, for instance, when the factors weighed by an algorithm do not correspond to any “real-world features”,<sup>19</sup> similar to the *algorithmically invented* discrimination grounds or consumer vulnerabilities I mentioned earlier. These do not map onto what discrimination or vulnerability triggers are for humans. It seems that in these cases, any information meaningful to a human would not adequately reflect the internal logic of the algorithm and therefore would not be useful, while the code itself might.



---

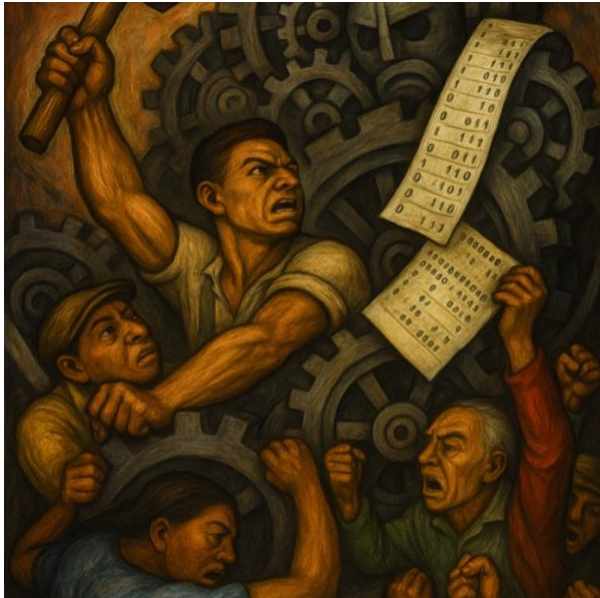
<sup>19</sup> Gianclaudio Malgieri, “Just” Algorithms: Justification (Beyond Explanation) of Automated Decisions Under the General Data Protection Regulation’ (2021) 1 Law and Business 16 <<https://doi.org/10.2478/law-2021-0003>>.

*In the second place*, the Court reduced the degree to which the syntactic process of algorithmic decision-making can be scrutinised and contested because it does not contain semantic meaning to a human.

The general understanding of the role of the right to explanation, also shared by the Court, is that it enables a subject of such a decision to contest it. The right to contest automated decisions is a guarantee that must be provided to the subject of an automated decision and is established in Art 22 (3) GDPR. The algorithm used in the decision-making *is* information pertaining to the logic involved in the decision-making, even if it is not meaningful information to a human. The syntactic information process, the algorithm itself, is certainly an essential part of the automated decision and must be open to contestation. For that, it must be open to scrutiny. Yet, the Court in *Dun & Bradstreet* effectively restricted the object of contestability to semantic information about that algorithm, a second-hand description of “the procedure and principles”, which could as well be – deliberately or not – incomplete or incorrect.

Granted, Alice or another average data subject on their own would probably not make heads or tails of the algorithm if it were provided. It is important that, where possible, such data subjects can understand at least some parts of the automated decisions they are subjected to.

However, this should not exclude the possibility of scrutinising the algorithm itself by tools other than explanation, when such scrutiny is desirable, and the necessary tools and expertise are available.



Think of the various forms of the algorithmic audit, possibly using other AI tools, that examine if bias or errors are built into the code, or if the algorithm relies on “appropriate mathematical or statistical procedures”,<sup>20</sup> meaning that the science behind the algorithm is correct. While Alice will likely not conduct such an audit, the digital rights groups could. And we cannot overstate how important civil society has been in enforcing the data protection rights and algorithmic accountability in recent years, and that using data subject information rights as a tool made this civil activism possible. To name a few examples, Worker Info Exchange and App Drivers & Couriers Union have been the driving force behind a successful challenge and a demand for an explanation of the Uber and Ola “robofiring” their drivers based on the algorithmic score.<sup>21</sup> Activists of the Panoptikon Foundation in Poland used their GDPR data access rights to expose and challenge the opacity of the behavioural advertising industry.<sup>22</sup> Limiting the contestability of algorithms to what an average data subject can understand pre-empts algorithmic auditing by civil society groups.



To be sure, it was the language of the GDPR that set up the Court for an impossible task to square the circle, to explain what information would help a human understand the internal logic of an algorithm.

---

<sup>20</sup> As Recital 71 GDPR prescribes

<sup>21</sup> <https://www.workerinfoexchange.org/post/historic-digital-rights-win-for-wie-and-the-adcu-over-uber-and-ola-at-amsterdam-court-of-appeal>

<sup>22</sup> [https://responsibledata.io/rd-reflection-stories/building-collective-momentum-to-challenge-the-ad-tech-industry/?utm\\_source=chatgpt.com](https://responsibledata.io/rd-reflection-stories/building-collective-momentum-to-challenge-the-ad-tech-industry/?utm_source=chatgpt.com)

## b Machine learning models as personal data

The disconnection between the algorithmic information processes and the human-centric assumptions about information will manifest in other parts of the GDPR.

**One** such point of disconnection is the status of machine learning models as personal data.



Models are *syntactic* information and do not carry semantic meaning to humans. For instance, Large Language Models are a combination of word fragments – or tokens - and their statistical relationships. The European Data Protection Board has recently concluded that under some circumstances, AI models are personal data. *But why is that?*

*One* possible line of argument is that the machine learning models, while not being semantic information themselves, *contain* semantic information from their training dataset, including personal data, but in a scrambled way.<sup>23</sup> A model contains personal data when the scramble can be unscrambled, and the personal data from the training dataset can be teased out of the model. But this argument has a very limited use. It gives the data protection rights to the people whose data was used to train models, and not those who are *affected* by the models and might benefit from the GDPR protections the most.

---

<sup>23</sup> Michael Veale, Reuben Binns and Lilian Edwards, 'Algorithms That Remember: Model Inversion Attacks and Data Protection Law' (2018) 376 *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 20180083 <<https://doi.org/10.1098/rsta.2018.0083>>.



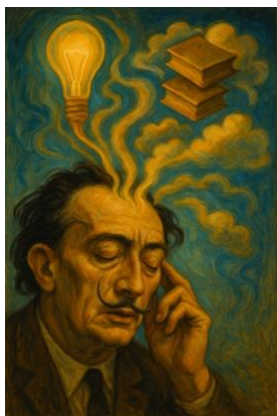
An alternative argument – which does not exclude the first one – is based on acknowledging that syntactic information can be personal data in its own right and thus subject to the GDPR guarantees. Personal data in the GDPR is defined as *any* information relating to an identified or identifiable natural person. According to the caselaw, this equally includes information *about* people, as well as information *not* about people if its processing is intended to or will likely impact people. This formulation is broad enough to include syntactic information such as algorithms, used in contexts where they produce effects on people’s lives, be it through automated decisions, unfair treatment, credit scoring and other circumstances. This argument means that the people who are affected by the models have rights regarding those models, resulting in much more useful protections.

The Board seems to be more inclined in favour of the first argument. Some national supervisory authorities, for instance, the Data Protection Authority of Hamburg, have refused to consider AI models personal data on the grounds that they “lack individual information content”, and so they might be ready to follow the former line of reasoning and not the latter.

The *trouble is* that this position goes against the established practice of considering other syntactic information as personal data. Most notably, tissue samples and genetic data on their own do not carry any meaning to humans. Yet, the European Court of Human Rights in *Marper* explicitly considered human tissue samples as personal data under the Data Protection Convention No 108. The GDPR also explicitly recognises genetic data as sensitive personal data that deserves special protection. *Therefore*, refusing to count AI models as personal data on the ground of a lack of semantic meaning would undermine the integrity and internal consistency of the data protection law.

## 5 A broader legal problem

As algorithms proliferate and cause disputes, other domains of law will, without a doubt, face similar tensions. I have already mentioned that discrimination and consumer law are facing a crisis. They both rely on the outdated assumptions that unfair treatment is a human cognitive process based on traits constructed and understood by humans, such as race, gender, disability and age.



Freedom of expression was conceived in the era of *human* expression and will likely face similar issues. Its scope and limitations will become relevant in the context of the regulation of AI-generated content. Think, for instance, about measures against AI-generated fake news, which could be considered as interference with expression.

Freedom of expression was not intended to cover synthetic speech. But the academic agreement - so far - is that the AI speech is protected under Article 10 of the European Convention of Human

Rights. AI does not have its own ideas to impart, but De Vries argues that AI content, for instance, images, jokes and other text, can be enjoyable and enriching to humans.<sup>24</sup> This is *their right to access* AI-generated information and “*to participate in ... meaning-making*” that is protected.

One discussion to be had here is to what extent AI-generated content should be protected as a source of human meaning-making when it is hard-wired to hallucinate, or in other words, present factually false information even when the science, the syntactic processes of learning and reproducing statistical relationships between data points, are correct. Large Language Models are also said to be hard-wired to lie and manipulate, skills they “mastered” by learning from patterns of human communication.

If the reason of this protection is to safeguard a human cognitive process of meaning making, should the same protection apply to the syntactic distortion of these semantic cognitive processes? Or, should the *integrity* of the human cognitive processes be protected against the syntactic intervention?

## 6 What can be done?

So, what can be done? Where should we look for answers?

---

<sup>24</sup> Michael Veale, Reuben Binns and Lilian Edwards, ‘Algorithms That Remember: Model Inversion Attacks and Data Protection Law’ (2018) 376 *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 20180083 <<https://doi.org/10.1098/rsta.2018.0083>>. 100.

## a Law & technology research

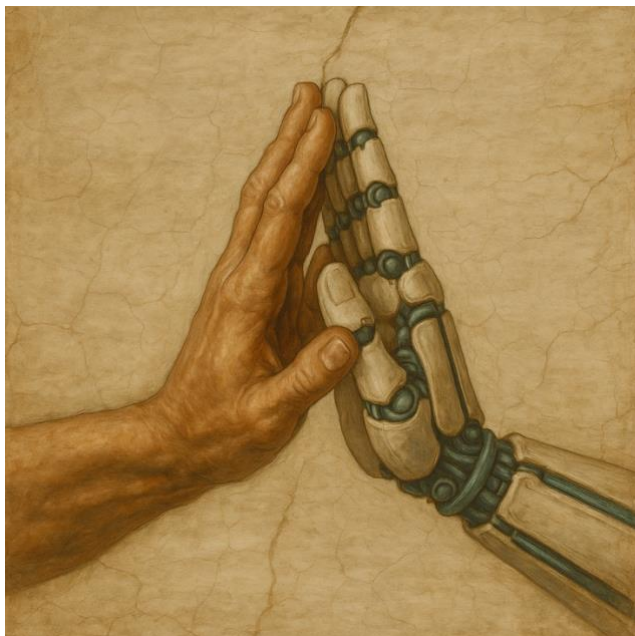
I see two ways to approach this mismatch between how the law *thinks* information works and how it *actually* works. These two ways coincidentally map onto how the research into law and technological change is organised at law departments.

The *first* way is to continue domain-specific efforts to map and address how this mismatch affects each isolated area of law. Consumer law scholars at the *private* law departments will continue looking for solutions to the problem of algorithmic consumer vulnerability. Equality, administrative and human rights scholars at the *public* law departments will continue proposing alternative ways to capture and control algorithmic discrimination, for instance by proposing new categories that should be considered as protected, or /--/ how algorithmic decisions by public authorities can be explained. These are all, without a doubt, valid lines of research. The *trouble* is that such a setup encourages the researcher to limit their gaze to the boundaries of that isolated legal domain. In these settings, researchers rarely talk across domains, while they face essentially the same issues.

The *second* way, which I believe is much better, is for the researchers to talk to each other, to look beyond their isolated domains to see a bigger picture, uncover and solve broader legal problems related to technology. The law itself usually does not provide all the answers to all the questions, and so this kind of research looks to other disciplines to learn more about the nature and consequences of socio-technical change. This includes, but is not limited to, computer science, AI, *Science and Technology Studies*, critical data and media studies. This kind of multidimensional and interdisciplinary research agenda is usually carried out in a law and technology department.

Not every law school *must* have a *Law and Technology department*. But the choice to have one says something about what kind of research that Law School wants. I am grateful to the Utrecht School of Law for seeing value and choosing for this research, and giving us an opportunity to set up the “Data and AI research platform” for cross-domain interdisciplinary research into broader legal problems linked to data and AI.

## b Scientifically grounded understanding of information for law



From that cross-domain vantage point, there is a broader legal issue that I cannot unsee: the law as a whole must move beyond its casual approach to how it views and regulates information, especially in the algorithmic context. We need to develop an understanding of information for law, that is based on scientific study rather than colloquial use, and that would be equally suitable to regulate humans and information machines.

The law will regulate information phenomena, old and new, including algorithmic decision-making, synthetic information, discrimination or unfair consumer practices more effectively if grounded in a scientific understanding of what information is and how it impacts the world.

Some argue that defining information in law is neither necessary (but I hope I convinced you otherwise) nor possible.

I propose that the law adopt the *Pragmatic Approach to Information* as a starting point for developing a Pragmatic Approach to Information *in Law*. The Pragmatic Approach already helps scholars of information from other disciplines to talk to each other in a productive way, past the irreconcilable differences of whether information must have meaning to a human or if machines understand.

To recap, the Pragmatic Approach to Information *in Law* should acknowledge that information can equally exist on semantic and syntactic planes, that these planes are different and do not intersect, but the existence of one does not exclude the existence of the other.

There are **three ways** in which the Pragmatic Approach to Information *has value for law*.

**First**, it provides a *language* for legal rules to talk about information, equally useful in the human and algorithmic context.

**Second**, it gives us a *way to think about legal problems* involving information as either human meaning-driven, or syntactic, || meaning-agnostic. And based on that, it helps us choose regulatory responses. The art for the law would be to learn how to navigate those planes and choose a right one for the specific legal situation.

**Finally**, it helps the law *ask new questions*. For instance, if and which semantic spaces for human meaning-making and creativity should be shielded from the intervention of AI? Should some fields of human relationships, poetry, civil society and science be reserved for human meaning-making and exclude machines acting as humans?

These points will be central themes both in my research and education for the coming years.

## 7 Conclusions: when the difference makes the difference



Algorithms are invading and reshaping nearly every domain of human activity. As algorithmic processes gradually become the matter of legal norms, knowing the difference between how machines and humans think, and recognising when that difference *matters*, will be crucial if the law is to regulate algorithms effectively.

A few years ago, we might have gotten away with applying information concepts intuitively. This lack of a scientific basis for how law treats "information" surfaced only occasionally. But the difficult questions about the nature and impact of information will not stay isolated to a few legal domains. We might be witnessing the birth of a full-on crisis that cuts across the entire legal landscape.

To solve this crisis, we urgently need an updated understanding of information in law. The Pragmatic Approach to Information in Law provides a good start.

To develop it requires looking to other disciplines like information studies and computer sciences, to explore the conceptual options they offer.

We'll *also* need to connect this understanding to fundamental questions in law, like the structure of legal rules and legal interpretation. This will help find the right balance between legal certainty and the open texture of the law, amid the noise and confusion of competing information concepts.

Some of this effort will be painful, and all of it -- as best things in science -- will be slow. But doing nothing is not an option. Acting as if the clothes are real does not make the emperor any more dressed.

## 8 About humans – het dankwoord



I would not be where I stand today without the support of many. I thank the Executive Board, the Board of the Faculty of Law, Economics, and Governance, and the School of Law for their trust and the many opportunities for growth and academic leadership.

I am grateful to Ivo, Anne, José, Janneke, Eddy and Manuel for the warm welcome and for helping me find my way in a new place.

Over the past three and a half years in Utrecht, I have met many colleagues working on technology and technology governance. I am grateful for your openness to have a conversation across the boundaries of disciplines and domains.

I'd like to thank my fellow teachers in the Law & Technology LLM for your collegiality and critical reflection on what the core of teaching Law & Technology should be about. I am also grateful to my students, whom I have the honour to teach Privacy & Data Protection law. You ask sharp questions and remind me to stay critical of the commonly accepted truths.

Catalina, you are one of the hardest-working and well-versed academics I've met. Thank you for being my law & technology sister-in-arms.

Thank you to my INFO-LEG team - Mara, Raphael, Sebastian, Gijs, and Bryce. You made swimming against the current of the data protection research much less lonely and more enjoyable.

The acknowledgements would not be complete without mentioning Tilburg Institute for Law, Technology and Society. TILT has formed me as an academic. I hope to be a carrier of its culture of intellectual and human generosity and camaraderie. Eleni, Esther, Linnet and others will remain the benchmark of academic quality and keep me on my toes.

Mireille, Ronald and Bert-Jaap, thank you for the intellectual stimulation of your scholarship and your mentorship. You will forever define a law and technology academic to me.

Dear Corien, you bring out the best in people. You taught me to trust my own ideas, especially when they go against the commonly accepted. I owe my academic wings to you.

I wouldn't be who I am without my parents, their work ethic and unconditional love. They both have inquisitive minds, and I hope I am *like* them in this way. My dad should be watching on the livestream. Папа, спасибо тебе и маме за все. Аплодисменты в конце будут для вас. My dad is an engineer who, at 79, still likes building things and is a proud patent-holder for a thing in a rocket, so the interest in machines must be running in the genes.

My mom was one of the strongest people I've known. She was well-read. I grew up watching her balance her career with being the most caring mother, and so I never felt I had to choose one of the two. She taught me that no intellectual prowess is worth anything without kindness.

And finally, Jasper and Sacha. Jas, you give me more than I can say. You are my best friend, cheerleader and the most wonderful father to our son. The two of you are my home, my sun and my inspiration. Without you, none of it would be possible or worthwhile.

I have said.